

OM of: US-09-784-340-2 to: GenEmbl:* out_format : pfs
Date: Aug 27, 2001 8:11 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL-frame+ p2n.pool -DEV-xlp
-O=/cgn2.1/USPTO.spool/US09784340/runat_27082001.123146.90/app-query.fasta.1.591
-DB=GenEmbl -QMT=fastcap -SUFFIX=olip2n.rge -GAPOP=4.500
-GAPEXT=0.050 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-FGAPOP=4.500 -FGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
-TRANS=human40.cdl -LIST=45 -DOCALLIGN=200 -THR_SCORE=quality
-THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTEXT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09784340@cgn1.1.5822 -NCPU=6
-ICPU=3 -LONGLOGS -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-784-340-2

Query length: 527

Database: GenEmbl:*

Database sequences: 1344157

Database length: -856060004

Search time (sec): 1747.330000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

Score list:

Sequence	Strid Orig	ZScore	Escore	Len	Documentation
gb_p15:AK025587	+	441.00	8307.81	0.0	2786 AK025587 Homo sapiens cDNA: FI
gb_p17:AC021146	+	239.00	4457.30	4.9e-240	212904 AC021146 Homo sapiens chrom
gb_p11:AC040087	+	174.00	3259.12	2.7e-173	1650 AX040087 Sequence 10 from Patg
gb_p17:AC021146	+	75.00	1354.12	3.4e-67	212904 AC021146 Homo sapiens chrom
gb_p15:AC011254	+	42.00	731.21	1.7e-32	169246 AC011254 Homo sapiens chrom
gb_p15:AC011254	+	41.00	742.80	3.9e-33	1584 AF184901 Mus musculus olfacto
gb_p15:AC011254	+	41.00	742.09	4.0e-33	1766 AF006054 Homo sapiens mRNA for
gb_p15:AC011254	+	40.00	722.93	5.0e-32	1831 X57565 Rattus norvegicus mRNA
gb_p15:AC011254	+	33.00	589.65	1.3e-24	2079 AJ005162 Homo sapiens mRNA for
gb_p15:AC011254	+	33.00	589.61	1.3e-24	2091 AF081793 Homo sapiens UDP-gluc
gb_p15:AC011254	+	33.00	589.60	1.3e-24	2093 Y00317 Human mRNA for liver mi
gb_p15:AC011254	+	33.00	589.56	1.3e-24	2107 U59209 Homo sapiens C19orf2
gb_p15:AC011254	+	33.00	572.83	1.1e-23	2130 AF135416 Homo sapiens UDP-gluc
gb_p15:AC011254	+	33.00	559.03	6.7e-23	226077 AC013296 Homo sapiens clone
gb_p15:AC011254	+	30.00	532.79	1.9e-21	2108 AF112112 Macaca fascicularis U
gb_p15:AC011254	+	29.00	519.46	1.1e-20	895 M35086 Rat UDP-glucuronosyltran
gb_p15:AC011254	+	29.00	514.34	1.1e-20	1661 M13506 Rat liver UDP-glucuron
gb_p15:AC011254	+	28.00	496.50	2.0e-19	1662 AF294901 Macaca mulatta UDP-gluc
gb_p15:AC011254	+	28.00	496.15	2.1e-19	1753 AF016310 Macaca fascicularis U
gb_p15:AC011254	+	28.00	495.78	2.2e-19	1855 U05428 Human 3.4-catechol est
gb_p15:AC011254	+	28.00	493.46	3.0e-19	2648 U91582 Macaca fascicularis UDP
gb_p15:AC011254	+	28.00	493.10	3.1e-19	2799 X63359 H. sapiens UGT2B10 mRNA
gb_p15:AC011254	+	28.00	467.59	8.3e-18	139015 AF025488 Homo sapiens chrom
gb_p15:AC011254	+	27.00	477.58	2.3e-18	1662 AF294902 Macaca mulatta UDP-gluc
gb_p15:AC011254	+	27.00	477.17	2.4e-18	1768 AF112113 Macaca fascicularis U
gb_p15:AC011254	+	27.00	453.77	4.9e-17	63705 AF135419 Homo sapiens cDNA: FI
gb_p15:AC011254	+	27.00	448.67	9.4e-17	139015 AC025488 Homo sapiens chrom
gb_p15:AC011254	+	24.00	420.13	3.6e-15	1844 U06273 Rattus norvegicus Sprag
gb_p15:AC011254	+	23.00	400.86	3.6e-14	1947 U06273 Rattus norvegicus UDP-g
gb_p15:AC011254	+	23.00	400.44	4.5e-14	2075 AF072223 Macaca fascicularis U
gb_p15:AC011254	+	23.00	400.32	4.6e-14	2090 U08854 Human UDP-glucuronosylt
gb_p15:AC011254	+	23.00	400.29	4.6e-14	2114 AF180322 Homo sapiens UDP-gluc
gb_p15:AC011254	+	23.00	371.26	1.9e-12	2123 U06641 Human UDP-glucuronosylt
gb_p15:AC011254	+	22.00	380.18	6.1e-13	2547 AF175221 Cavia porcellus UDP-g
gb_p15:AC011254	+	21.00	363.41	5.3e-12	1832 AF01061 Oryctolagus cuniculus U
gb_p15:AC011254	+	19.00	325.97	6.4e-10	1722 AF015492 Homo sapiens UDP-gluc
gb_p15:AC011254	+	19.00	325.52	6.8e-10	1846 U06274 Rattus norvegicus Sprag
gb_p15:AC011254	+	19.00	324.69	7.5e-10	2097 AF064200 Homo sapiens UDP-gluc

gb_cm:RABUGT2BC + 19.00 322.58 9.9e-10 2897 I L01083 Oryctolagus cuniculu
gb_hlg13:AC055794 + 19.00 297.32 2.5e-08 138665 I AC055794 Homo sapiens chr
gb_cm:RABUGT2B + 18.00 307.25 7.1e-09 1671 I L01082 Oryctolagus cuniculu
gb_r02:RNUDPT + 18.00 307.07 7.2e-09 1716 I Y00156 Rat mRNA for hepatic
gb_r02:RATGRT + 18.00 306.69 7.6e-09 1319 I M3109 Rat UDP-glucuronosyl
gb_r01:MMUDPCT + 17.00 287.63 8.7e-08 1858 I X06358 Mouse mRNA for UDP-g

seq_name: gb_p15:AK025587

seq_documentation_block: 2786 bp mRNA PRI 29-SEP-2000

LOCUS AK025587

DEFINITION Homo sapiens cDNA: FLJ21934 fls, clone HEP04364.

ACCESSION AK025587

VERSION AK025587.1 GI:10438147

KEYWORDS oligo capping; fls (full insert sequence).

SOURCE Homo sapiens hepatoma cell_line: HepG2 cDNA to mRNA, clone_11b: HEP

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (sites)

Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,

Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,

Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.

NEDO human cDNA sequencing project.

Unpublished (2000)

2 (bases 1 to 2786)

Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,

Shibahara, T., Tanaka, T. and Nakamura, Y.

Direct Submission

Submitted (29-AUG-2000) to the DDBJ/EMBL/Genbank databases. Sumio

Sugano, Institute of Medical Science, University of Tokyo.

Laboratory of Genome Structure Analysis, Human Genome Center;

Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan

(E-mail: cdna@lelms.u-tokyo.ac.jp, Tel: 81-3-5449-5286,

Fax: 81-3-5449-5416)

NEDO human cDNA sequencing project supported by Ministry of

International Trade and Industry of Japan: cDNA full insert

sequencing. Research Association for Biotechnology: cDNA library

construction, 5'- & 3'-end one pass sequencing: Department of

Virology and Human Genome Center, Institute of Medical Science,

University of Tokyo (partly supported by Science and Technology

Agency).

Location/Qualifiers

1. 2786

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_line="HepG2"

/clone="HEP04364"

/note="cloning vector pME18SFL3"

/note="unnamed protein product"

/codon_start=1

/db_xref="GI:10438148"

/translation="MPDRTENEFVDAALVAVPLGSLTWSYIKNDFVIRGLTK

MCSEFTYVQTLTKLQETNYDMLDPIVPCGDIAELTAVPVTLTIISVGNRER

SGCKLPAPLSYVPMPTGLDPTMLEKNSMLSVLFHFMVTDYDHYHEFYSKAL

GPPTLCFVGAELIWLRTYWDPEPQVQNFPEVGLHCKPKAPALPKEMENYVS

SESDGIVLGSGLFONVTEKANKIATASLAQIPKVLVGGKPKSTGANTRLYD

IFQNDLVGPKTKAPRTAGNGACIYEAITHYCPVPCVTFGGQDINIMAKKGAAY

INFKTWSDELRLALRTVYDSSYENAKRLSRHHDPVKRLDRAVFTIEYMKHG

AHLRSAAIDLWTFQHSYIDVIGFLACVAIVAFLETKCFSCCKFNKTRIERE"

BASE COUNT 919 a 463 c 496 g 908 t

alignment_scores: quality: 441.00 length: 441

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment block:
US-09-784-340-2 x AK025587

Align seg 1/1 to: AK025587 from: 1 to: 2786

```

56 LeuThrHisSerLysProSerLeuIleAspTyrArgLysProSerAlaLeu 72
|||||
2 TTGACTGCTCAAGCCCTTCGTTAATTGACTACAGAAAGCCCTTCGCAATT 51
|||||
72 uLysPheGluValAlaHisMetProGlnAspArgThrGluGluAsnGlu 89
|||||
52 GAATATTGAGGTGGTCCATATCCACAGACAGAAACAGAAAGAAATGAAA 101
|||||
89 LePheValAspLeuAlaLeuAsnValLeuProGlyLeuSerThrTrpGln 105
|||||
102 TATTGTGACCTAGCTCTGAAATGCTTGGCCAGGCTTATCAACCTGGCAA 151
|||||
106 SerValIleLysLeuAsnAspPhePheValGluIleArgGlyThrLeuLys 122
|||||
152 TCAGTTAATAAATTAAATGATTTTTTTGTGAAATAGAGAACTTTAAA 201
|||||
122 MetMetCysGluSerPheIleTyrAsnGluThrLeuMetLysLysLeuG 139
|||||
202 AATGATGTGTGAGACTTTATCTACAAATCAGACGCTTATGAAGAAGCTAC 251
|||||
139 IngLufhrAsnTyrAspValMetLeuIleAspProValIleProCysGly 155
|||||
252 AGAAGAACCACTACGATGTAATGCTTATAGACCCGTGATCCCTGTGGA 301
|||||
156 AspLeuMetAlaGluLeuLeuAlaValProPheValLeuThrLeuArgIle 172
|||||
302 GACCTGATGGCTGAGGCTTGCACAGCTCCCTTTGTGCTCACACTTGAAT 351
|||||
172 eSerValGlyLysAsnMetGluArgSerCysGlyLysLeuProIleArg 189
|||||
352 TTCTGTAGAGAGCAATATGAGACGAGCTGTGGAACTTCCAGCTCCAG 401
|||||
189 eUserTyrValProValProMetThrGlyLeuThrAspArgMetThrPhe 205
|||||
402 TTTCCTATGTACCTGTGCTTATGACAGACTAACAGACAGATGACCTTT 451
|||||
206 LeuGluArgValLysAsnSerMetLeuSerValLeuPheHisPheTrpIle 222
|||||
452 CTGGAAGAGATAAATAAATCAATCTTTCAGCTTTGTCCACTCTGCAAT 501
|||||
222 eGlnAspTyrAspTyrHisPheTrpGluGluPheTyrSerLysAlaLeuG 239
|||||
502 TCAGGATTACACATCATTTTGGGAAAGATTATATAGTAAGCATTTAG 551
|||||
239 LysArgProThrThrLeuCysGluThrValGlyLysAlaGluIleTrpLeu 255
|||||
552 GAAGGCCACATACATATATGTGAGACTGTGGAAAGCTGAGATATGGCTA 601
|||||
256 IleArgThrTyrTrpAspPheGluPheProGlnProTyrGlnProAsnPh 272
|||||
602 ATACGAACATATGGAATTTGAATTTCTCAACCATACCAACCTAATCTT 651
|||||
272 eGluPheValGlyLysLeuHisCysLysProAlaLysAlaLeuProLysG 289
|||||
652 TGAGTTGTTGGAGGATTGCCACTTAACCTGCCAAAGCTTTGCCCTAAG 701
|||||
289 LysMetGluAsnPheValGlnSerSerGlyLysAspGlyIleValAlaPhe 305
|||||
702 AAATGAAAAATTTTGTCCAGATTCAGAGGAGATGATATTGGTGCTTT 751
|||||
306 SerLeuGlySerLeuPheGlnAsnValThrGluGluLysAlaAsnIleIle 322
|||||
752 TCCTCTGGGGCTACTGTTCAAAATGTTACAGAAAGAAAGGCTAATATCAT 801
|||||
322 eAlaSerAlaLeuAlaGlnIleProGlnLysValLeuThrPArgTyrLysG 339
|||||
802 TGCTTCAGCCCTTCCAGATCCACAGAAAGGTATATGAGGTACCAAG 851
|||||

```

```

339 LysLysProSerThrLeuGlyAlaAsnThrArgLeuTyrAspTrpIle 355
|||||
852 GAAAAAACCATCCATTCATTAGAGCCCAATCTCGGTGATATGGATA 901
|||||
356 ProGlnAsnAspLeuLeuGlyHisProLysThrLysAlaPheIleThrIle 372
|||||
902 CCCAGAAATGATCTTCTTGTCATCCCAAAACCAAGCTTTATATCACTCA 951
|||||
372 eGlyLysMetAsnGlyIleTyrGluAlaIleTyrHisGlyValProMetLys 389
|||||
952 TGGTGGAATGAATGGATCTATGAAAGCTATTACCATGGGGTCCCTATGG 1001
|||||
389 aGlyValProIlePheGlyAspGlnLeuAspAsnIleAlaHisMetLys 405
|||||
1002 TGGGATTCCTCATATTGGTGTATCAGCTTGATTAACATAGCTTACATGAAG 1051
|||||
406 AlaLysGlyAlaAlaValGluIleAsnPheLysThrMetThrSerGluAs 422
|||||
1052 GCCAAAGGAGCAGCTGTGAAATAAATCACTTCAAACTATGACAGGGAAGA 1101
|||||
422 PheLeuArgAlaLeuArgThrValIleThrAspSerSerTyrLysGluA 439
|||||
1102 TTACTGAGGGCTTGAAGACAGTCATACCGATTCCTCTTATAAAGAGA 1151
|||||
439 snAlaMetArgLeuSerArgIleHisAspGlnProValLysProLeu 455
|||||
1152 ATGCTATAGATTATCAAGAATTCACCATGATCAACCTGTAAAGCCCTTA 1201
|||||
456 AspArgAlaValPheTrpIleGluPheValMetArgHisLysGlyAlaLys 472
|||||
1202 GATCGAGCAGCTTCTGTGATGAGTTTGTCAATGCCCAAGAGAGCCAA 1251
|||||
472 SHisLeuArgSerAlaAlaHisAspLeuThrTrpPheGlnHisTyrSerI 489
|||||
1252 GCACCTGGATCAGCTGCCATGACCTCAGCTGTTCCAGCAGCACTACTTA 1301
|||||
489 LeaspValIleGlyPheLeuLeu 496
|||||
1302 TAGATGTGATTGGGTTCTCTGCTG 1324
|||||

seq_name: gb_hhg7:AC021146
seq_documentation block:
LOCUS: AC021146 212904 bp DNA HTG 07-JUL-2000
DEFINITION: Homo sapiens chromosome 4 clone RFL1-468NI4, WORKING DRAFT
SEQUENCE: 24 unchromed pieces.
ACCESSION: AC021146
VERSION: AC021146.4 GI:8568861
KEYWORDS: HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE: human.
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE: 1 (bases 1 to 212904)
AUTHORS: Waterston, R.H.
TITLE: The sequence of Homo sapiens clone
JOURNAL: Unpublished
REFERENCE: 2 (bases 1 to 212904)
AUTHORS: Waterston, R.H.
TITLE: Direct Submission
JOURNAL: Submitted (14-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108 USA
COMMENT: On Jun 16, 2000 this sequence version replaced gi:7344259.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0468NI4
----- Summary Statistics -----

```

Sequencing vector: M13; 88%
Chemistry: Dye-primer ET; 88% of reads
Chemistry: Dye-terminator Big Dye; 12% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 199729 bases at least Q40
Consensus quality: 203731 bases at least Q30
Consensus quality: 206340 bases at least Q20
Insert size: 213000; agarose-fp
Insert size: 210604; sum-of-ctdigs
Quality coverage: 3.75 in Q20 bases; agarose-fp
Quality coverage: 3.83 in Q20 bases; sum-of-ctdigs

NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1564: contig of 1564 bp in length
1565 1664: gap of unknown length
1665 3085: contig of 1421 bp in length
3086 3185: gap of unknown length
3186 5134: contig of 1949 bp in length
5135 5234: gap of unknown length
5235 8169: contig of 2935 bp in length
8170 8269: gap of unknown length
8270 10341: contig of 2072 bp in length
10342 10441: gap of unknown length
10442 13614: contig of 3173 bp in length
13615 13714: gap of unknown length
13715 17089: contig of 3375 bp in length
17090 17189: gap of unknown length
17190 20701: contig of 3512 bp in length
20702 20801: gap of unknown length
20802 25001: contig of 4200 bp in length
25002 25101: gap of unknown length
25102 29020: contig of 3919 bp in length
29021 29120: gap of unknown length
29121 33356: contig of 4236 bp in length
33357 33456: gap of unknown length
33457 36902: contig of 3446 bp in length
36903 37002: gap of unknown length
37003 42975: contig of 5973 bp in length
42976 43075: gap of unknown length
43076 48339: contig of 5264 bp in length
48340 48439: gap of unknown length
48440 56529: contig of 8090 bp in length
56530 56629: gap of unknown length
56630 65213: contig of 8584 bp in length
65214 65313: gap of unknown length
65314 74715: contig of 9402 bp in length
74716 74815: gap of unknown length
74816 88546: contig of 13731 bp in length
88547 88646: gap of unknown length
88647 103367: contig of 14721 bp in length
103368 103467: gap of unknown length
103468 118167: contig of 14700 bp in length
118168 118267: gap of unknown length
118268 132765: contig of 1498 bp in length
132766 132865: gap of unknown length
132866 154092: contig of 21227 bp in length
154093 154192: gap of unknown length
154193 182144: contig of 27952 bp in length
182145 182244: gap of unknown length
182245 212904: contig of 30660 bp in length

Location/Qualifiers
1..212904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"

BASE COUNT 69110 a 36915 c 37432 g 67121 t 2326 others
ORIGIN
alignment_scores:
Quality: 239.00 Length: 239
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-784-340-2 x AC021146/rev ..
Align seg 1/1 to reverse of: AC021146 from: 1 to: 212904
1 MetArgSerSplySerAlaLeuValPheLeuLeuLeuLeuLeuPheCy 17
87413 ATGAGGCTGACAGAGTGCAGCTTGCTGCTGCTGCTGCTGCTG 87364
17 sValGlyCyGlyPheCyGlyValLeuValTrrProCyAspMetS 34
87363 TGTGGCTGTGGATCTGTGGAAAGTCGTGGTGGCCCTGTGACATGA 87314
34 eRHSrPleuAsnValValIleLeuGluGluLeuValArgGly 50
87313 GCCATTGGCTTAATGTCAAGCATTTCTAGAGAGCTCATAGTAGAGGC 87264
51 HisGluValThrValLeuThrHisSerLysProSerLeuLeuAspYrrAc 67
87263 CATGAGTAACTAGTATGACTCAACGCTTCGTTAATTGACTACAG 87214
67 GlyProSerAlaLeuLysPheGluValValHisMetProGluAspArgT 84
87213 GAAGCTTCGTGCAATTGAATTGAGTGTCATPAGCCACAGACAGAA 87164
84 hrgLgluAsnGluIlePheValAspLeuAlaLeuAsnValLeuProGly 100
87163 CAGAGAAATGAATATTGTGACCTAGCTGATGATCTGTGCCAGGC 87114
101 LeuSerThrTrpGlnSerValIleLysLeuAsnAspPheValGluI 117
87113 TTATCAACCTGGCAATCAGTTAATAAATGATTTTGTGTGAAT 87064
117 eArgGlyThrLeuLysMetCysGluSerPheIleTyrAsnGluThrL 134
87063 AAGAGCACTTTAAAAATGATGTGTGAGACCTTTATCTACATCAGACGC 87014
134 eumetLysLysLeuGluGluThrAsnTyrAspValMetLeuLeuAspPro 150
87013 TTATGAAGAGCTACAGAAACCACTAGCATGATATGCTTATAGACCT 86964
151 ValIleProCysGlyAspLeuMetAlaGluLeuLeuAlaValProPheVa 167
86963 GTGATTCCTCCGTGAGACCTGATGGCTGATGCTGCAATCCCTTTGCT 86914
167 lleuThrLeuArgIleSerValIleGlyLysMetGluArgS-CyGlyL 184
86913 GCTTCACACTTAGAATTTCTGTAGAGCAATATGAGCCAGACCTGTGGCA 86864
184 yslLeuProAlaProLeuSerTyrValProValProMetThrGlyLeuThr 200
86863 AACCTTCAGCTCCACTTCTCTAGTACCTGTGCTATGAGAGGACATACA 86814
201 AspArgMetThrPheLeuGluValArgValLysAsnSerMetLeuSerValIle 217
86813 GACAGATGACCTTCTCGAAGAGTAAATAATCAATGCTTTCAGAGTTT 86764
217 uphHisPheTrpIleGlnAspTyrAspTyrHisPheTrpGluIuphet 234
86763 GTTCACACTTCTGATTCAGGATTCAGCATTCATTTTGGGAAGAGTTT 86714
234 yrrSerLysAlaLeuGly 239
|||||

86713 ATAGTAGCATTAAGCT 86697

seq_name: gb_pat1:AX040087

seq_documentation_block:

LOCUS AX040087 1650 bp DNA PAT 18-NOV-2000

DEFINITION Sequence 10 from Patent WO0063351.

ACCESSION AX040087

VERSION AX040087.1 GI:112330049

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1650)

AUTHORS Lal, P., Yue, H., Tang, Y. T., Hillman, J. L., Baughn, M. R. and Yang, J.

TITLE Carbohydrate-modifying enzymes

JOURNAL Patent: WO 0063351-A 10 26-OCT-2000;

Incyte Genomics, Inc. (US)

FEATURES

Source 1.1650

/organism="Homo sapiens"

/db_xref="taxon:9606"

/note="Incyte ID No: 2912330CBI"

BASE COUNT 489 a 330 c 354 g 477 t

ORIGIN

alignment_scores:

Quality: 174.00 Length: 174

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-784-340-2 x AX040087

Align seg 1/1 to: AX040087 from: 1 to: 1650

```

354 TrrllleProGlnAsnAspleuGlyHisProlystHrlyslAaPhel1 370
|||||
1100 TGGATACCCAGATGATCTTCTGGTCATCCCAAAACCAAGCTTTAT 1149
|||||
370 ePrrHISGLyMeTasnglylletYrguAlaileYrHISGLyValP 387
|||||
1150 CATCTATGGGGAATGATGGATCTATGAGCTATTTCACATGGGGTCC 1199
|||||
387 rOMeValGlyValProIlePheGlyAspGlnLeuAspAsnIleAlaHis 403
|||||
1200 CTATGGTGGAGTCCCATATTGGTATGATGAGCTGATAACATAGCTCAC 1249
|||||
404 MetIlyAlaIysGlyAlaAlaValGluIleAsnPhelystHrMetHrse 420
|||||
1250 ATGAAGGCCAAAGAGACAGCTGTAAGAAATTAACCTCAAAACTATGACAA 1299
|||||
420 rGluAspleuLeuArgAlaLeuArgThrValIleThrAspSerSerTyrL 437
|||||
1300 CGAAGATTTCAGAGGCTTGGAGAACAGCATATACCATCTCTTATA 1349
|||||
437 ySGLuAsnAlaMetArgLeuSerArgIleHisAspGlnProValLys 453
|||||
1350 AAGAGATGCTATGAGATTCATCAAGAAATTCACCATGATCAACTGTAAAG 1399
|||||
454 ProLeuAspArgAlaValPheTrpIleGluPheValMetArgHISLysG 470
|||||
1400 CCCCTGATGACAGCAGCTCTTCTGATGAGTGTGTCAGCCACAAAG 1449
|||||
470 yAlaIysHISLeuArgSerAlaAlaHISAspLeuThrTrpPheGlnHIS 487
|||||
1450 AGCCAGACACTGCGATGACCTGCCATGACCTGCTTCCAGACT 1499
|||||
487 ySerIleAspValIleGlyPheLeuLeuThrCysValAlaThrAlaIle 503
|||||
1500 ACTCTATAGATGATGAGTGGTCTCTGCTGACCTGTGCGCAACTGCTATA 1549

```

504 PheLeuPheThrLysCysPheLeuPheSerCysGlnLysPheAsnLysrh 520
 1550 TTTCTGTTCACAAAATGTTTTTATTTCTCTGTCACAAAATTTAATAAAC 1599

520 rArgLysIleGluLysArgGlu 527
 1600 TAGAAAGATAGAAAGAGCGAA 1621

seq_name: gb_htg7:AC021146

seq_documentation_block:

LOCUS AC021146 212904 bp DNA HTG 07-JUL-2000

DEFINITION Homo sapiens chromosome 4 clone RP11-468N14, WORKING DRAFT

SEQUENCE, 24 unordered pieces.

ACCESSION AC021146

VERSION AC021146.4 GI:8568861

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 212904)

AUTHORS Waterston, R. H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 212904)

AUTHORS Waterston, R. H.

TITLE Direct Submission

JOURNAL Submitted (14-JAN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT On Jun 16, 2000 this sequence version replaced gi:7344259.

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0468N14
----- Summary Statistics -----
Sequencing vector: M13; 88%
Sequencing vector: plasmid; 12%
Chemistry: Dye-Primer ET; 88% of reads
Chemistry: Dye-terminator Big Dye; 12% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 199729 bases at least Q40
Consensus quality: 203731 bases at least Q30
Consensus quality: 206340 bases at least Q20
Insert size: 210604; agarose-fp
Insert size: 210604; sum-of-contigs
Quality coverage: 3.75 in Q20 bases; pharose-fp
Quality coverage: 3.83 in Q20 bases; sum-of-contigs
----- NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1564: contig of 1564 bp in length
* 1565 1664: gap of unknown length
* 1665 3085: contig of 1421 bp in length
* 3086 3185: gap of unknown length
* 3186 5134: contig of 1949 bp in length
* 5135 5234: gap of unknown length
* 5235 8169: contig of 2935 bp in length
* 8170 8269: gap of unknown length
* 8270 10341: contig of 2072 bp in length
* 10342 10441: gap of unknown length
* 10442 13614: contig of 3173 bp in length

```

```

* 13615 13714: gap of unknown length
* 13715 17089: contig of 3375 bp in length
* 17090 17189: gap of unknown length
* 17190 20701: contig of 3512 bp in length
* 20702 20801: gap of unknown length
* 20802 25001: contig of 4200 bp in length
* 25002 25101: gap of unknown length
* 25102 29020: contig of 3919 bp in length
* 29021 29121: gap of unknown length
* 29121 33356: contig of 4236 bp in length
* 33357 33456: gap of unknown length
* 33457 36902: contig of 3446 bp in length
* 36903 42975: gap of unknown length
* 42976 43075: gap of unknown length
* 43076 48339: contig of 5264 bp in length
* 48340 56529: gap of unknown length
* 56530 56629: gap of unknown length
* 56630 65213: contig of 8584 bp in length
* 65214 65313: gap of unknown length
* 65314 74716: contig of 9402 bp in length
* 74716 74815: gap of unknown length
* 74816 88546: contig of 13731 bp in length
* 88547 103367: gap of unknown length
* 103368 103467: contig of 14721 bp in length
* 103468 118167: gap of unknown length
* 118168 118267: gap of unknown length
* 118268 132765: contig of 14498 bp in length
* 132766 132865: gap of unknown length
* 132866 154092: contig of 21227 bp in length
* 154093 154192: gap of unknown length
* 154193 182144: contig of 27952 bp in length
* 182145 182244: gap of unknown length
* 182245 212904: contig of 30660 bp in length
* 182245 212904: Location/Qualifiers
  source          1. 212904
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="RP11-468N14"
                  /chromosome="4"

```

```

BASE COUNT      69110 a 36915 c 37432 g 67121 t 2326 others
ORIGIN

```

```

alignment_scores:
  quality: 75.00      Length: 75
  Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-784-340-2 x AC021146 ..

Align seg 1/1 to: AC021146 from: 1 to: 212904

```

361 LeuGlyHisProLysThrLysAlaPheIleThrsIsglyMetAsnG1 377
|||||
33906 CTAGGTATCCCAAAACCAAGCTTTATCATCTGATGGAATGAGTGG 33955
|||||
377 YIleTyvGlaAlaIleTyvRhsGlyValPrometValGlyValProIleP 394
|||||
33956 GATCTATGAAGCTATTACCATGGGGTCCCTATGCTGGAGTTCCTCATAT 34005
|||||
394 heGlyAspGlnLeuAspAsnIleAlaHisMetLysAlaLysGlyAlaAla 410
|||||
34006 TTGGTGATGAGCTTGATACATACATGACATGAAGGCAAGAGCAGCT 34055
|||||
411 ValGluIleAsnPhelYsThrMetThSergLusAspLeuArgAlaIle 427
|||||
34056 GTAGAAATTAACCTCAAAACTATGACAAAGCAAGATTACTGAGGCTTT 34105
|||||
427 UArgThValIleThrsAspSer 435

```

```

34106 GAGAAAGTCAATACGATTCCTCG 34130
seq_name: gb_h33:AC011254

```

seq_documentation_block:

```

LOCUS      AC011254 169246 bp DNA HTG 26-MAY-2000
DEFINITION Homo sapiens chromosome 4 clone RP11-401E5 map 4, WORKING DRAFT
SEQUENCE, 12 unordered pieces.
AC011254

```

AC011254.3 GI:7107765

VERSION AC011254.3

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

```

*****
NOTE: This is a 'working draft' sequence. It currently
consists of 12 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
*****
1 2121: contig of 2121 bp in length
2 2122 2221: gap of 100 bp
2222 5198: contig of 2977 bp in length
*****

```

434 SerSerTyrLysGluAsnAlaMetArgLeuSerArgIleHisAspG1 450
|||||
16824 TCTAGTTATTAAGAGATGCTATGAGCTTATCAAGATTCCACCATGATCA 16873

435 serTLVSGI¹AS²AL³AL⁴me⁵tr⁶g⁷leu⁸ser⁹tr¹⁰g¹¹leu¹²SH¹³AS¹⁴GI¹⁵Pr 451
1303 TCTTATTAAGAGAAATGCCATGCAGACTATCAAGATCCACCATGACACACC 1352
451 ovalLVSP¹ro²leu³aas⁴p⁵tr⁶g⁷al⁸val⁹l¹⁰ph¹¹er¹²tr¹³l¹⁴eg¹⁵l¹⁶u¹⁷p¹⁸he¹⁹val²⁰me²¹tr²²g²³h 468
1353 AGTGAAGCCCTCGGACCGACAGCTCTTCGAGATTAGATTGTCAATGCGCTC 1402

468 IslysglyAlalysHsleuarg 475
 1403 ACAAAGAGCAAGACCTTCCT 1425

seq_name: gb_pr9: HSA6054

seq_documentation_block:

LOCUS HSA6054 1766 bp mRNA PRI 16-JUL-1999
 DEFINITION Homo sapiens mRNA for UDP glucuronosyltransferase.
 ACCESSION AJ006054
 VERSION AJ006054.1 GI:4753765
 KEYWORDS UDP glucuronosyltransferase; ugt2A1 gene.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1766)
 Jedititschky G., Cassidy A.J., Sales M., Pratt N. and Burchell B.

TITLE Cloning and characterization of a novel human olfactory

JOURNAL UDP-glucuronosyltransferase
 Biochem. J. 340 (Pt 3), 837-843 (1999)

REFERENCE 2 (bases 1 to 1766)
 Cassidy A.J.

AUTHORS Direct Submission

JOURNAL Submitted (07-MAY-1998) Cassidy A.J., Molecular and Cellular
 Pathology, University of Dundee, University Dept. of Molecular and
 Cellular pathology, Ninewells Hospital, Dundee, DD1 9SY, SCOTLAND

FEATURES Location/Qualifiers

1..1766

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="olfactory"

64..1647

/gene="ugt2A1"

64..1647

/FC_number="2.4.1.17"

/codon_start=1

/product="UDP glucuronosyltransferase"

/protein_id="CAB41974.1"

/db_xref="GI:4753766"

/translation="MLNNLLLSIQLSLIGTTLGNYLIMPGSHMLNKKIITDELL
 KKHNTVIVASGALFTTPSPSLFEIYKVPFGKRTGCVIKDPSTMLENRPSPS
 TIRFYQEMAKVIEEFVLSRGICDGLKNEKMTKQKGFELVSDVPCGDIVA
 LKLGIPFVSLRSPASTVEKHCQKVPSPYVPAVLSLTDMSFDRVRNISTRM
 ODWMEFTLKMSPSYSKALGRPTLCEIEMGKAEIMLRITMDPEEPRLPNEFVG
 GLCKKPAKPLPKMEEFVORTSGEYGVVFLSGVMKMLEKANLISALAOIPKYL
 WRYKGPATLGSNTRFLPDMIPONDLGKHTKRAFTHGCTNGIYEAITYGIPWCV
 MFADQPNIAHMAKGAIVNMTMTSADLSAVRAVINEPPTKENAMLSRIHDDQ
 PVKPLDRAVFIEMFVMRKGAHKLRYAANDLSFQYHSLDVIQFLVCVTTALFLVIO
 CCLFSCOKFKIGTKKKRRE"

BASE COUNT 542 a 359 c 363 g 502 t

ORIGIN

alignment_scores:

Quality: 41.00 Length: 41

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block: 1

US-09-784-340-2 x HSA6054

Align seg 1/1 to: HSA6054 from: 1 to: 1766

435 SerTYLYSGluAsnAlaMetArgLeuSerArgIleHisHisAspGlnIrr 451
 1366 TCTTAAAGAGAAATGCTATGAGTTATACAGATTCACCATGATGATAC 1415
 1416 TGTAAAGCCCTGATGAGAGCTCTTGATGAGATTGTCATCGGCC 1465

468 IslysglyAlalysHsleuarg 475
 1466 ACAAAGAGCAAGACCTTCGG 1488

seq_name: gb_r02: RRUGTGENE

seq_documentation_block:

LOCUS RRUGTGENE 1831 bp DNA ROD 14-SEP-1999
 DEFINITION Rattus norvegicus mRNA UDP-glucuronosyltransferase 2A1 precursor.
 ACCESSION X57565
 VERSION X57565.1 GI:57762
 KEYWORDS UDP-glucuronosyltransferase; ugt2A1 gene.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 1831)
 Lazard D., Zupko K., Porta Y., Nef P., Lazarovits J., Horn S.,
 Khien M. and Lancel D.

TITLE Odorant signal termination by olfactory UDP glucuronosyl

JOURNAL transferase
 Nature 349 (6312), 790-793 (1991)

REFERENCE 2 (bases 1 to 1831)
 Cassidy A.J.

AUTHORS Direct Submission

JOURNAL Submitted (07-MAY-1998) Cassidy A.J., Molecular and Cellular
 Pathology, University of Dundee, University Dept. of Molecular and
 Cellular pathology, Ninewells Hospital, Dundee, DD1 9SY, SCOTLAND

FEATURES Location/Qualifiers

1..1831

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

64..1647

/gene="ugt2A1"

64..1647

/FC_number="2.4.1.17"

/codon_start=1

/product="UDP-glucuronosyltransferase 2A1 precursor"

/protein_id="CAA40797.1"

/db_xref="GI:3980217"

/translation="MLKNLLLSIQLSLIGTTLGNYLIMPGSHMLNKKIITDELL
 KKHNTVIVASGALFTTPSPSLFEIYKVPFGKRTGCVIKDPSTMLENRPSPS
 TIRFYQEMAKVIEEFVLSRGICDGLKNEKMTKQKGFELVSDVPCGDIVA
 LKLGIPFVSLRSPASTVEKHCQKVPSPYVPAVLSLTDMSFDRVRNISTRM
 ODWMEFTLKMSPSYSKALGRPTLCEIEMGKAEIMLRITMDPEEPRLPNEFVG
 GLCKKPAKPLPKMEEFVORTSGEYGVVFLSGVMKMLEKANLISALAOIPKYL
 WRYKGPATLGSNTRFLPDMIPONDLGKHTKRAFTHGCTNGIYEAITYGIPWCV
 MFADQPNIAHMAKGAIVNMTMTSADLSAVRAVINEPPTKENAMLSRIHDDQ
 PVKPLDRAVFIEMFVMRKGAHKLRYAANDLSFQYHSLDVIQFLVCVTTALFLVIO
 CCLFSCOKFKIGTKKKRRE"

BASE COUNT 505 a 393 c 432 g 500 t 1 others

ORIGIN

alignment_scores:

Quality: 40.00 Length: 40

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block: 1

US-09-784-340-2 x RRUGTGENE

Align seg 1/1 to: RRUGTGENE from: 1 to: 1831

436 TYLYSGluAsnAlaMetArgLeuSerArgIleHisHisAspGlnIrr 452
 1369 TATTAAGAGAAATGCTATGAGTTATACAGATTCACCATGATGATAC 1418
 1419 GAAGCCCTGATGAGAGCTCTTGATGAGATTGTCATGCGCACA 1468
 469 ysglyAlalysHsleuarg 475
 1469 AAGGAGCCCAAGCACCTTCCT 1488

```

seq_name: gb_pt9:HSNJ5162
seq_documentation_block: 2079 bp mRNA PRI 24-JUN-1998
LOCUS HSNJ5162
DEFINITION Homo sapiens mRNA for UDP-glucuronosyltransferase.
ACCESSION AJ005162
VERSION AJ005162.1 GI:3135024
KEYWORDS UDP-glucuronosyltransferase; UGT2B4 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Ritter,J.K.
TITLE 1 (bases 1 to 2079)
JOURNAL Submitted (08-MAY-1998) Ritter J.K., Pharmacology and Toxicology,
of Virginia Commonwealth University, P.O. Box 980613, Medical College
of Virginia, Richmond, Virginia, 23112-0613, USA
2 (bases 1 to 2079)
Ritter,J.K., Chen,F., Sheen,Y.Y., Lubet,R.A. and Owens,I.S.
Two human liver cDNAs encode UDP-glucuronosyltransferases with 2
log differences in activity toward parallel substrates including
hydroxycholeic acid and certain estrogen derivatives
Biochemistry 31 (13), 3409-3414 (1992)
92207964
3 (bases 1 to 2079)
Jackson,M.R., McCarthy,L.R., Harding,D., Wilson,S., Coughtrie,M.W.
and Burchell,B.
Cloning of a human liver microsomal UDP-glucuronosyltransferase
cDNA
Biochem. J. 242 (2), 581-588 (1987)
87241362
FEATURES
source
location/Qualifiers
1..2079
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Lambda bacteriophage gt11"
/tissue_lib="Liver"
34..1620
/gene="UGT2B4"
34..1620
/gene="UGT2B4"
/EC_number="2.4.1.17"
/function="transfers glucuronic acid to hydroxycholeic
acid and estrogen catechols"
/codon_start=1
/evidence-experimental
/product="UDP-glucuronosyltransferase"
/protein_id="CAA06396.1"
/db_xref="GI:3135025"
/db_xref="SPTREMBL:O60867"
/translation="MSKWTSAALLIQSCYFSSGCKVLWPIESHMNIKTIID
ELVQKHEVTVLASASTISDPNSPSTLKEFYVPTLSTEFEDIIKLVKRMALPK
DTFMSYEQVOEIMWTENDILRKFCDIVSKKMKKLOESRFDDVLAIVPFEELP
AELKIPFVYSLRSPGYAIKESGGLFPPSYVAVVSELSDDWTFLEPVKMTIYL
YEEFMFOLEDMKKMDQFSEVIGRPTTILSETMAKADIVLINYMDFOFPHLLPVEE
VGGIHCRAKRLPKMEEFNOSGNGVYVSLGMSVNTSEERANVIAKLAKIPK
VLMRPDKKPTDGLNTRLYKWTIPNDLLGHPKTRAFTHGGANCIYAIHGIPIWG
VPLRADPDNTAHMKAGAAVSLDFHTSSDLDNALNTVINDPYKRNAMKLSIHH
DQPKPLDRAVFIETFVIRHKAGAKHLRVAADLTWFOYHSLDVIGFLACVATVFIIL
TKLFCVWKFPVTKGKKGRD"
546
/conflict
/gene="UGT2B4"
/citation=[3]
/replace="g"
548..549
/gene="UGT2B4"
/citation=[3]
/replace="ct"
1187..1188
/gene="UGT2B4"
/conflict

```

```

seq_name: gb_pt4:AF081793
seq_documentation_block: 2091 bp mRNA PRI 18-AUG-1998
LOCUS AF081793
DEFINITION Homo sapiens UDP glucuronosyltransferase 2B4 precursor (UGT2B4)
mRNA, UGT2B4*1109, L396 variant allele, complete cds.
ACCESSION AF081793
VERSION AF081793.1 GI:3426331
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Jin,C.J., Miners,J.O., Lillywhite,K.J. and Mackenzie,P.I.
TITLE 1 (bases 1 to 2091)
JOURNAL J Clin. Invest. 105, 1111-1117 (2000)
MEDLINE 93326164
REFERENCE
AUTHORS Mackenzie,P.I.
TITLE 2 (bases 1 to 2091)
JOURNAL Direct Submission
Submitted (30-JUL-1998) Clinical Pharmacology, Flinders University
of South Australia, Bedford Park, SA 5042, Australia
FEATURES
source
location/Qualifiers
1..2091
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
1..2091
/gene="UGT2B4"
/allele="Formerly UGT2B11"
/allele="UGT2B4*1109, L396 variant"
36..1622
/gene="UGT2B4"
/conflict

```

```

/citation=[3]
/replace="ct"
1190
/gene="UGT2B4"
/citation=[3]
/replace="c"
1192
/gene="UGT2B4"
/citation=[3]
/replace="a"
1977..1978
/citation=[3]
/replace="cga"
2068
/citation=[3]
/replace="g"
BASE COUNT 634 a 396 c 436 g 613 t
ORIGIN
alignment_scores:
Quality: 33.00 Length: 33
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-784-340-2 x HSNJ5162
Align seq 1/1 to: HSNJ5162 from: 1 to: 2079
443 LeuSerArgIleHisHisAspGlnProValIysProLeuAspArgAlaVala 459
|||||
1366 TTAATCAGAAATTCATCATGATCAACGATGAAGCCCTTATCATGACAGT 1415
459 lphETpRlEGluPhEvaImetargHisLysGlyAlaLysHisLeuArg 475
|||||
1416 CTTCGTGATTGAATTGTCTCATGCGCCATTAAGGAGCCACACCTTCGG 1464

```



```
/codon_start=1
/product="UDP-glucuronosyltransferase 2B4 precursor"
/db_xref="GI:3426332"
/translation="MSKMTSALLLIQLSCYSSGCGKVLVPTFESHMMNKTLLD
ELVORHEVTVLASSASISFDPSNSTLKEVYVSLTKTEFEDIIKOLVKRAEELPK
DIFWYSLOVOELMTFNDILRKFCDIVSNKMLKLOESRDVVLADAVPEFGELL
AELKIPDYSLRPFSPGYALIEKHSGLLFPSPYVPMSELSDQMTFIERVKMIVL
YEFEMFOIPMKKMOFSEVIGRPTTLSEPMKADIMLIRNYWDOFPHLLPNVEF
VGLHCKRPKPLPKKKEFYOSSGNGVYVFSLSGWSVNTSEERANVIALAKTIPOK
VLMRFGNKPDTGLNTRLYKWIIPONDLSHPKTRPAFTTHGANGIYEALYH1PMVG
VPLFADQPNIAHMAKKAQAVSLDFHTMSSTDLLNALKTVINDPLVKNMKSRIIHH
DQVVKLDRAVVEIEFVMRHKGAHLRVAHADLTPWQYHSLDVTGFLACVAVTIFII
TKCLFCVMKFEVFRGKKGRD"
BASE COUNT      638 a      401 c      437 g      615 t
ORIGIN

alignment_scores:
    Quality:      33.00      Length:      33
    Ratio:        1.000      Gaps:        0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x AF081793 ..
Align seg 1/1 to: AF081793 from: 1 to: 2091

443 LeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaIva 459
|||||
1368 TTATCAGAAATTCATCATGATCAACAGCAGCCCTTGATCGAGCAGT 1417

459 lPheTrpIleGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475
|||||
1418 CTCTGGATTGAATTTGTCATGCGCATTAAGGACCAAGCACCTTCGG 1466

seq_name: gb_pr10:HSUDPCT
seq_documentation_block:
LOCUS      HSUDPCT      2093 bp      mRNA      PRI      23-MAR-1995
DEFINITION Human mRNA for liver microsomal UDP-glucuronosyltransferase
              (UDPCT).
ACCESSION   Y00317
VERSION     Y00317.1 GI:37588
KEYWORDS    transferase; UDP-glucuronosyltransferase.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 2093)
AUTHORS    Burchell,B.
TITLE      Direct Submission
JOURNAL    Submitted (26-JAN-1987) Burchell B., Dept. Biochemistry, Medical
          Sciences Institute, University Dundee, Scotland
AUTHORS    Jackson,M.R., McCarthy,L.R., Harding,D., Wilson,S., Coughtrie,M.W.
          and Burchell,B.
TITLE      Cloning of a human liver microsomal UDP-glucuronosyltransferase
          cDNA
JOURNAL    Biochem. J. 242 (2), 581-588 (1987)
MEDLINE    87241362
FEATURES
    source
        location/Qualifiers
            1..2093
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /tissue_type="liver"
            /clone_lib="lambda gtl1 liver cDNA"
            38..106
            /product="put. signal peptide"
            38..1624
            /note="precursor"
            /codon_start=1
```

```
/protein_id="CAA68415.1"
/db_xref="GI:37589"
/db_xref="SWISS-PROT:P06133"
/translation="MSKMTSALLLIQLSCYSSGCGKVLVPTFESHMMNKTLLD
ELVORHEVTVLASSASISFDPSNSTLKEVYVSLTKTEFEDIIKOLVKRAEELPK
DIFWYSLOVOELMTFNDILRKFCDIVSNKMLKLOESRDVVLADAVPEFGELL
AELKIPDYSLRPFSPGYALIEKHSGLLFPSPYVPMSELSDQMTFIERVKMIVL
YEFEMFOIPMKKMOFSEVIGRPTTLSEPMKADIMLIRNYWDOFPHLLPNVEF
VGLHCKRPKPLPKKKEFYOSSGNGVYVFSLSGWSVNTSEERANVIALAKTIPOK
VLMRFGNKPDTGLNTRLYKWIIPONDLSHPKTRPAFTTHGANGIYEALYH1PMVG
VPLFADQPNIAHMAKKAQAVSLDFHTMSSTDLLNALKTVINDPLVKNMKSRIIHH
DQVVKLDRAVVEIEFVMRHKGAHLRVAHADLTPWQYHSLDVTGFLACVAVTIFII
TKCLFCVMKFEVFRGKKGRD"
BASE COUNT      638 a      397 c      439 g      619 t
ORIGIN

mat_peptide
misc_feature
misc_feature
misc_feature
BASE COUNT      638 a      397 c      439 g      619 t
ORIGIN

alignment_scores:
    Quality:      33.00      Length:      33
    Ratio:        1.000      Gaps:        0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x HSUDPCT ..
Align seg 1/1 to: HSUDPCT from: 1 to: 2093

443 LeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaIva 459
|||||
1370 TTATCAGAAATTCATCATGATCAACAGCAGCCCTTGATCGAGCAGT 1419

459 lPheTrpIleGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475
|||||
1420 CTCTGGATTGAATTTGTCATGCGCATTAAGGACCAAGCACCTTCGG 1468

seq_name: gb_pr10:HSU59209
seq_documentation_block:
LOCUS      HSU59209      2107 bp      mRNA      PRI      02-JUL-1998
DEFINITION Homo sapiens C19steroid specific UDP-glucuronosyltransferase mRNA,
              complete cds.
ACCESSION   U59209
VERSION     U59209.1 GI:3287472
KEYWORDS    UGT2B17G.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 2107)
AUTHORS    Beaulieu,M., Levesque,E., Hum,D.W. and Belanger,A.
TITLE      Isolation and characterization of a novel cDNA encoding a human
          UDP-glucuronosyltransferase active on C19 steroids
JOURNAL    J. Biol. Chem. 271 (37), 22855-22862 (1996)
MEDLINE    96394358
REFERENCE   2 (bases 1 to 2107)
AUTHORS    Hum,D.W., Belanger,A., Beaulieu,M. and Levesque,E.
TITLE      Direct Submission
JOURNAL    Submitted (23-MAY-1996) Laboratory of Molecular Endocrinology,
          Centre Hospitalier de l'universite Laval, 2705 Boul. Laurier,
          Ste-Foy, Quebec G1V 4G2, Canada
JOURNAL    1..2107
FEATURES
    source
        location/Qualifiers
            1..2107
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /tissue_type="prostate"
            /cell_type="LNCaP"
            52..1644
            /codon_start=1
```

```

/note="UGT2B17"
/codon_start=1
/product="C19steroid specific UDP-glucuronosyltransferase"
/protein_id="AAC3549.1"
/db_xref="GI:3287473"
/translation="MSLKMWSVFLMLQSCYFSSGCGKYLWPFTEFSHNMNKTILE
ELVORGEHIVLSSASILVNASKSAIKLEVPTSLTKNDLEDFMKMDRMYSIS
KNFWSYFQLOELCMESYDNIKLCEDAVLNMKMLKLOESKEDVLADAVNCGEL
LAELINIPFLYSLEFSEVYTERKNGGFLPPSPVPMSELSDOMTPEMRKMTYM
LYEDFMFOADLKKMDQFSEVLGRPTLFTETMCKAEMMTLIRYMDPEFRPLPND
FVGLHCKPAKPLPKMEPEFVOSGNGIYFVSGISMSSESNAMISALAQIDQ
KYLMDRKKRPNLTGSTRILKMLPQNDLGHPTKAFITHGNGIYEAIGHIPIY
GILFADQHDNIAMKAKGALSVDITMSSRDILNALKSVINDPIYKEMIKLSRTH
HDQVPLDRVAVMIEFVMRHKGAKHLRVAHNLWIOYHSLDVIATLACVATMIRM
ITKCLFCFPRKLAKTGKKRRD"
BASE COUNT      657 a      382 c      424 g      644 t
ORIGIN
alignment_scores:
  Quality:      33.00      Length:      33
  Ratio:        1.000      Gaps:        0
Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-784-340-2 x HSU59209
Align seg 1/1 to: HSU59209 from: 1 to: 2107
443 LeuserArgIIeHISAspGlnProValIysProLeuAspArgAlaVala 459
|||||
1367 TTATCAAGAAATTCATCATGATCAACCGGTGAAGCCCTTGATCGAGCAGT 1436
459 IPhetrIleGluPheValMetArgHISLysGlyValAlaIysHISLeuArg 475
|||||
1437 CTTCGGATTGATTGTTCATCGCCATTAAGGACCAACGACCTTCGG 1485
seq_name: gb-pr5:AF135416
seq_documentation_block:
LOCUS      AF135416      27310 bp      DNA
DEFINITION Homo sapiens UDP-glucuronosyltransferase gene, complete cds.
ACCESSION  AF135416
VERSION    AF135416.1 GI:8650277
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 27310)
AUTHORS   Riedy,M. and Miller,A.
TITLE     Genomic organization and structure of the UGT2B gene complex at
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 27310)
AUTHORS   Riedy,M. and Miller,A.
TITLE     Direct Submission
JOURNAL   Submitted (18-MAR-1999) Axyx Pharmaceuticals, 11099 N Torrey Pines
          Rd., La Jolla, CA 92039, USA
FEATURES
source     location/Qualifiers
            1..27310
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="4"
            /map="4q13"
            /clone="Research Genetics BAC 321d14"
            complement(join<1169..1445,5738..5957,7139..7226,
            9699..10100,14223..14371,15671..>163911)
            /product="UDP-glucuronosyltransferase"
            complement(join(1169..1445,5738..5957,7139..7226,
            9699..10100,14223..14371,15671..163911)
            /codon_start=1
            /product="UDP-glucuronosyltransferase"

```

```

/protein_id="AAF78145.1"
/db_xref="GI:8650278"
/translation="MSMKWTSAALLIOLSCYFSSGCGKYLWPFTEFSHNMNKTILD
ELVORGEHIVLSSASISFDPNPSSTLKEEYVSLTKNEFEDIIKOLVRAMELKP
DTEWSYFQLOELCMESYDNIKLCEDAVLNMKMLKLOESKEDVLADAVNCGEL
LAELINIPFLYSLEFSEVYTERKNGGFLPPSPVPMSELSDOMTPEMRKMTYM
LYEDFMFOADLKKMDQFSEVLGRPTLFTETMCKAEMMTLIRYMDPEFRPLPND
FVGLHCKPAKPLPKMEPEFVOSGNGIYFVSGISMSSESNAMISALAQIDQ
KYLMDRKKRPNLTGSTRILKMLPQNDLGHPTKAFITHGNGIYEAIGHIPIY
GILFADQHDNIAMKAKGALSVDITMSSRDILNALKSVINDPIYKEMIKLSRTH
HDQVPLDRVAVMIEFVMRHKGAKHLRVAHNLWIOYHSLDVIATLACVATMIRM
ITKCLFCFPRKLAKTGKKRRD"
BASE COUNT      8660 a      5083 c      4901 g      8466 t
ORIGIN
alignment_scores:
  Quality:      33.00      Length:      33
  Ratio:        1.000      Gaps:        0
Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-784-340-2 x AF135416/rev
Align seg 1/1 to reverse of: AF135416 from: 1 to: 27310
443 LeuserArgIIeHISAspGlnProValIysProLeuAspArgAlaVala 459
|||||
1423 TTATCAAGAAATTCATCATGATCAACCGGTGAAGCCCTTGATCGAGCAGT 1374
459 IPhetrIleGluPheValMetArgHISLysGlyValAlaIysHISLeuArg 475
|||||
1373 CTTCGGATTGATTGTTCATCGCCATTAAGGACCAACGACCTTCGG 1325
seq_name: gb.htg4:AC013296
seq_documentation_block:
LOCUS      AC013296      226077 bp      DNA
DEFINITION Homo sapiens clone RP11-3M18, LOW-PASS SEQUENCE SAMPLING.
ACCESSION  AC013296
VERSION    AC013296.5 GI:9121251
KEYWORDS   HTG; HTGS; PHASE0.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 226077)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens, clone RP11-3M18
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 226077)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhallal,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Donnell,D., Doyle,M.,
            Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczy,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Melidrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Testfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Jul 13, 2000 this sequence version replaced gi:6514003.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L2775
Center clone name: 3_M18

* NOTE: This record contains 246 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 814: contig of 814 bp in length
815 914: gap of 100 bp
915 1704: contig of 790 bp in length
1705 1804: gap of 100 bp
1805 2602: contig of 798 bp in length
2603 2702: gap of 100 bp
2703 3484: contig of 782 bp in length
3485 3584: gap of 100 bp
3585 4378: contig of 794 bp in length
4379 4478: gap of 100 bp
4479 5257: contig of 779 bp in length
5258 5357: gap of 100 bp
5358 6150: contig of 793 bp in length
6151 6250: gap of 100 bp
6251 7038: contig of 788 bp in length
7039 7138: gap of 100 bp
7139 7940: contig of 802 bp in length
7941 8040: gap of 100 bp
8041 8869: contig of 829 bp in length
8870 8969: gap of 100 bp
8970 9790: contig of 821 bp in length
9791 9890: gap of 100 bp
9891 10695: contig of 805 bp in length
10696 10795: gap of 100 bp
10796 11592: contig of 797 bp in length
11593 11692: gap of 100 bp
11693 12475: contig of 783 bp in length
12476 12575: gap of 100 bp
12576 13364: contig of 789 bp in length
13365 13464: gap of 100 bp
13465 14255: contig of 791 bp in length
14256 14355: gap of 100 bp
14356 15142: contig of 787 bp in length
15143 15242: gap of 100 bp
15243 16023: contig of 781 bp in length
16024 16123: gap of 100 bp
16124 16932: contig of 809 bp in length
16933 17032: gap of 100 bp
17033 17878: contig of 846 bp in length
17879 17978: gap of 100 bp
17979 18786: contig of 808 bp in length
18787 18886: gap of 100 bp
18887 19677: contig of 791 bp in length
19678 19777: gap of 100 bp
19778 20575: contig of 798 bp in length
20576 20675: gap of 100 bp
20676 21463: contig of 788 bp in length
21464 21563: gap of 100 bp
21564 22350: contig of 787 bp in length
22351 22450: gap of 100 bp
22451 23220: contig of 770 bp in length
23221 23320: gap of 100 bp
23321 24108: contig of 788 bp in length
24109 24208: gap of 100 bp

24209 24969: contig of 761 bp in length
24970 25069: gap of 100 bp
25070 25852: contig of 783 bp in length
25853 25952: gap of 100 bp
25953 26726: contig of 774 bp in length
26727 26826: gap of 100 bp
26827 27625: contig of 799 bp in length
27626 27725: gap of 100 bp
27726 28540: contig of 815 bp in length
28541 28640: gap of 100 bp
28641 29413: contig of 773 bp in length
29414 29513: gap of 100 bp
29514 30314: contig of 801 bp in length
30315 30414: gap of 100 bp
30415 31215: contig of 801 bp in length
31216 31315: gap of 100 bp
31316 32099: contig of 784 bp in length
32100 32199: gap of 100 bp
32200 32995: contig of 796 bp in length
32996 33095: gap of 100 bp
33096 33892: contig of 797 bp in length
33893 33992: gap of 100 bp
33993 34782: contig of 790 bp in length
34783 34882: gap of 100 bp
34883 35658: contig of 776 bp in length
35659 35758: gap of 100 bp
35759 36554: contig of 796 bp in length
36555 36654: gap of 100 bp
36655 37443: contig of 789 bp in length
37444 37543: gap of 100 bp
37544 38361: contig of 818 bp in length
38362 38461: gap of 100 bp
38462 39228: contig of 767 bp in length
39229 39328: gap of 100 bp
39329 40157: contig of 829 bp in length
40158 40257: gap of 100 bp
40258 41059: contig of 802 bp in length
41060 41159: gap of 100 bp
41160 41959: contig of 800 bp in length
41960 42059: gap of 100 bp
42060 42830: contig of 771 bp in length
42831 42930: gap of 100 bp
42931 43733: contig of 803 bp in length
43734 43833: gap of 100 bp
43834 44657: contig of 824 bp in length
44658 44757: gap of 100 bp
44758 45554: contig of 797 bp in length
45555 45654: gap of 100 bp
45655 46440: contig of 786 bp in length
46441 46540: gap of 100 bp
46541 47356: contig of 816 bp in length
47357 47456: gap of 100 bp
47457 48260: contig of 804 bp in length
48261 48360: gap of 100 bp
48361 49165: contig of 805 bp in length
49166 49265: gap of 100 bp
49266 50058: contig of 793 bp in length
50059 50158: gap of 100 bp
50159 50968: contig of 810 bp in length
50969 51068: gap of 100 bp
51069 51841: contig of 773 bp in length
51842 51941: gap of 100 bp
51942 52748: contig of 807 bp in length
52749 52848: gap of 100 bp
52849 53670: contig of 822 bp in length
53671 53770: gap of 100 bp
53771 54554: contig of 784 bp in length
54555 54654: gap of 100 bp
54655 55442: contig of 788 bp in length
55443 55542: gap of 100 bp
55543 56341: contig of 799 bp in length
56342 56441: gap of 100 bp
56442 57261: contig of 820 bp in length

```

* 57262 57361: gap of 100 bp
* 57362 58170: contig of 809 bp in length
* 58171 58270: gap of 100 bp
* 58271 59048: contig of 778 bp in length
* 59049 59148: gap of 100 bp
* 59149 59954: contig of 806 bp in length
* 59955 60054: gap of 100 bp
* 60055 60858: contig of 804 bp in length
* 60859 60958: gap of 100 bp
* 60959 61761: contig of 803 bp in length
* 61762 61861: gap of 100 bp
* 61862 62631: contig of 770 bp in length
* 62632 62731: gap of 100 bp
* 62732 63524: contig of 793 bp in length
* 63525 63624: gap of 100 bp
* 63625 64411: contig of 787 bp in length
* 64412 64511: gap of 100 bp
* 64512 65269: contig of 758 bp in length

```

alignment_scores:

```

Quality: 33.00 Length: 33
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-784-340-2 x AC013296 ..

Align seg 1/1 to: AC013296 from: 1 to: 226077

443 LeuserATGTTGHSASPGInProvallyspProleuAsparGalaVala 459

178261 TTATCAAGAAATTCATCATGATCAACCAAGCCCTTGATCGACAGCAT 178310

459 lPheTrrpIeGlupheValmetArgHsIysGlyAlaIysHisIeuArg 475

178311 CTCTGATTCATATTGTGATGCGCCATTAAGAGCCCAAGCACCCTTCGG 178359

seq_name: gb-prs:AF112112

seq_documentation_block:

```

LOCUS AF112112 2108 bp mRNA PRI 11-APR-1999
DEFINITION Macaca fascicularis UDP-glucuronosyltransferase 2B19 precursor,
mRNA, complete cds.
ACCESSION AF112112
VERSION AF112112.1 GI:4580601
KEYWORDS
SOURCE
ORGANISM
. crab-eating macaque.
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
AUTHORS Belanger,G., Barbier,O., Hum,D.W. and Belanger,A.
TITLE Molecular cloning, expression and characterization of a monkey
steroid UDP-glucuronosyltransferase, UGT2B19, that conjugates
testosterone
JOURNAL Eur. J. Biochem. 260 (3), 701-708 (1999)
MEDLINE 99203465
REFERENCE
AUTHORS Belanger,G., Barbier,O., Hum,D.W. and Belanger,A.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-1998) Molecular Endocrinology, CHUL Research
Center, 2705 Laurier Blvd, Sainte Foy, Quebec G1V 4G2, Canada
FEATURES
SOURCE
1..2108
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
26..1612
/EC_number="2.4.1.17"
/note="UGT2B19"
/codon_start=1
/product="UDP-glucuronosyltransferase 2B19 precursor"
/protein_id="AAD24435.1"

```

```

/db_xref="GI:4580602"
/translation="MSMKWTSAALLIQLSCYLSFGSGKVLVMPTEFSHMNIKITLD
ELVQRHEVTVLAVSTSLPDPNNPSPLEKICPTSLTEFQDSVTOLVRRSDIRK
DTFWPHFLHVOEMMTYGDIMRRECKDVSNKTKMKKLOESRQDVLAOLAISPCGELL
AELKLPFYVSLRSPGYALEKHGGFLPPSPVPTVMTSELRDQMTFMRVOMATYV
YRPFMFQVNDVKRMVDQYSKVLGRPTTLEFNAAKATILIRNWDQFPHPLPNTVEF
VGLRHCKPAKPLPEKMEEFVQSSGDGVVVFSLGSVSNSEERANVIALAKTIQK
VLMRFQGNKPDITGLNTOLYKWLPPQNDLGHPRTRAFITHGANGIYEA1YHGIPWG
VPLFADQPNIAIMKAKAAVRLDPDMSSTDLNALKTVINDPIYKENAMKLSIHH
DQVFKELDRAVENIIEFVNRHKGAKHLRVAHDLTFQYHSLDIVIGFLACVATVIFII
TKCLFCVMKFEVTRKKGRD"
BASE COUNT 649 a 393 c 438 g 628 t
ORIGIN

```

alignment_scores:

```

Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-784-340-2 x AF112112 ..

Align seg 1/1 to: AF112112 from: 1 to: 2108

446 lLeHSHSAspgInProvallyspProleuAsparGalaValaPheTrrpI 462

1367 ATTCAATCATGATCAACCAAGCCCTTGATCGACAGATTTCTGGAT 1416

462 eGlupheValmetArgHsIysGlyAlaIysHisIeuArg 475

1417 TGAATTGTCTATGCGCCATTAAGAGCCCAACACCTTCGG 1456

THIS PAGE BLANK (USPTO)